OIPE
FEB 13 2006

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Patent Application

5

Applicant(s):	Cooper et al.
Docket No.:	YOR920000753US1
Serial No.:	09/785,650
Filing Date:	February 16, 2001
Group:	2655
Examiner:	Minerva Rivero

10

I hereby certify that this paper is being deposited on this date with the U.S. Postal Service as first class mail addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Signature: _____ Date: February 10, 2006

Title: Tracking Time Using Portable Recorders and Speech Recognition

15

APPEAL BRIEF

Mail Stop Appeal Brief - Patents
20 Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

25

Applicants hereby appeal the final rejection dated October 31, 2005, of claims 1-8, 10-22, and 24-30 of the above-identified patent application.

REAL PARTY IN INTEREST

30   The present application is assigned to International Business Machines Corporation, as evidenced by an assignment recorded on May 18, 2001 in the United States Patent and Trademark Office at Reel 011819, Frame 0975. The assignee, International Business Machines Corporation, is the real party in interest.

35   ## RELATED APPEALS AND INTERFERENCES
There are no related appeals or interferences.

STATUS OF CLAIMS

The present application was filed on February 16, 2001 with claims 1-28, of which claims 1, 15, 19, 24, 27, and 28 are independent claims. Claims 9 and 23 were cancelled in the Amendment and Response to Office Action dated March 16, 2004.

5   Claims 29 and 30 were added in the Amendment and Response to Office Action dated August 25, 2004. Claims 1-8, 10-22, and 24-30 are currently pending in the above-identified patent application. Claims 1-7 and 27 remain rejected under 35 USC §103(a) as being unpatentable over Logan et al., U.S. Patent No. 5,721,827, hereinafter "Logan," in view of Yahagi et al., U.S. Patent No. 4,984,274, hereinafter "Yahagi." Claims 8, 10-

10   22, 24-26, and 28-30 remain rejected under 35 USC §103(a) as being unpatentable over Logan in view of Yahagi, and further in view of Ladd et al., U.S. Patent No. 6,539,359, hereinafter "Ladd."

STATUS OF AMENDMENTS

15   There have been no amendments filed subsequent to the final rejection.

SUMMARY OF CLAIMED SUBJECT MATTER

The present invention is directed to the conversion speech, preferably recorded on a portable recorder, to text, analyzes the text, and determines voice

20   commands and times when the voice commands occurred. (Page 5, line 5, to page 10, line 14.) Task names are associated with voice commands and time segments. (Page 10, lines 15-27.) These time segments and tasks may be packaged as time increments and stored (e.g., in a file or database) for further processing. (Page 13, line 17, to page 19, line 14.) Preferably, phrase grammar rules are used when analyzing the text, as this helps

25   to determine voice commands. Using phrase grammar rules also allows the text to contain a variety of topics, only some of which are pertinent to tracking time. (Page 22, line 6, to page 23, line 20.)

30

STATEMENT OF GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Claims 1-7 and 27 are rejected under 35 USC §103(a) as being unpatentable over Logan, in view of Yahagi, and claims 8, 10-22, 24-26, and 28-30 are rejected under 35 USC §103(a) as being unpatentable over Logan in view of Yahagi, and

5   further in view of Ladd.

ARGUMENT

Independent Claims 1, 15, 19, 24, 27 and 28

Independent claims 1 and 27 are rejected under 35 USC §103(a) as being

10   unpatentable over Logan in view of Yahagi, and claims 15, 19, 24, and 28 are rejected under 35 USC §103(a) as being unpatentable over Logan in view of Yahagi, and further in view of Ladd.   Regarding claim 1, the Examiner acknowledges that Logan lacks "determining a first and second time associated with a speaking of a first and second of the voice commands, wherein said first and second voice command identifies a start and

15   end of said time interval and storing data identifying said time interval and data identifying one or more of said first voice command and second voice command," but asserts that Yahagi discloses these limitations.

Appellants note that Yahagi is directed to a speech recognition apparatus for *controlling a stop watch* with means for preventing errors due to delay in the speech

20   recognition process (see, Abstract), and that Logan is directed to an *audio program and message distribution system* in which a host system organizes and transmits program segments to client subscriber locations (see, Abstract).   Appellants maintain that Yahagi and Logan are therefore directed to unrelated art and that there is no motivation to combine the cited references.   Appellants have also found no disclosure or suggestion in

25   either Yahagi or Logan to combine the features cited by the Examiner.   Thus, a person of ordinary skill in the art would not look to combine Yahagi and Logan.

Regarding the Examiner's assertion that Yahagi discloses storing data identifying said time interval (stores measurement time data) and data identifying one or more of said first voice command and second voice command (col. 3, line 33, to col. 4,

30   line 14), Appellants note that the Examiner asserts that Yahagi teaches to *store*

3

measurement time data. Measurement time data, however, does *not "identify"* a time interval. For example, two time intervals may have the same measurement time data, but have different identities.

In addition, in the text cited by the Examiner, Yahagi teaches that "it

5   *stores measurement time data* from the measuring circuit 14 in *an address area designated from a plurality of data memory areas* 17a, 17b, . . . , of a RAM 17 and designates the next memory area." (Col. 3, lines 39-42; emphasis added.) Appellants could find no disclosure or suggestion by Yahagi to *store data identifying one or more of a first voice command and second voice command.*

10   Thus, contrary to the Examiner's assertion, Appellants note that neither Logan nor Yahagi disclose or suggest *storing data identifying a time interval and data identifying one or more of the first voice and second voice commands.* Independent claims 1, 19, and 27 require storing *data identifying said time interval* and *data identifying one or more of said first voice command and second voice command.*

15   Independent claims 15, 24, and 28 require determining a plurality of time increments, *each time increment comprising one of the tasks and at least one of the times*; and *storing one or more of said time increments* (wherein each task is associated with at least one of the times and at least one of the voice commands).

Thus, Logan and Yahagi, alone or in combination, do not disclose or

20   suggest storing data identifying said time interval and data identifying one or more of said first voice command and second voice command, as required by claims 1, 19, and 27, and do not disclose or suggest determining a plurality of time increments, each time increment comprising one of the tasks and at least one of the times; and storing one or more of said time increments, as required by claims 15, 24, and 28.

25   Additional Cited References

Ladd was also cited by the Examiner for its disclosure of converting each of at least two voice commands to text. Appellants note that Ladd is directed to a markup language for interactive services (see, Abstract). Ladd does *not* address the issue of determining and storing time intervals/increments, or storing data identifying voice

30   commands.

Thus, Ladd does not disclose or suggest storing data identifying said time interval and data identifying one or more of said first voice command and second voice command, as required by claims 1, 19, and 27, and does not disclose or suggest determining a plurality of time increments, each time increment comprising one of the tasks and at least one of the times; and storing one or more of said time increments, as required by claims 15, 24, and 28.

Claim 2

Claim 2 is rejected under 35 USC §103(a) as being unpatentable over Logan, in view of Yahagi. Regarding claim 2, the Examiner asserts that Logan discloses wherein the second voice command (BACK command subdivided into two commands) is implied because a predetermined time from the first voice command (predetermined time of segment) elapses before another voice command occurs (column 14, lines 11-27).

Appellants note that, in the text cited by the Examiner, Logan teaches resetting "the playback point to the beginning of the prior segment or the beginning of the prior subject description." (Col. 14, lines 15-17.) While the playback point may be reset, Logan does *not* disclose or suggest that the BACK command is executed or implied. Claim 2 requires wherein the second voice command is implied because a predetermined time from the first voice command elapses before another voice command occurs or because a predetermined ending time occurs and there is no voice command after the first voice command but before the predetermined ending time.

Thus, Logan, Yahagi, and Ladd, alone or in any combination, do not disclose or suggest wherein the second voice command is implied because a predetermined time from the first voice command elapses before another voice command occurs or because a predetermined ending time occurs and there is no voice command after the first voice command but before the predetermined ending time, as required by claim 2.

Claims 3, 4, 6, 17 and 20

Claims 3, 4, and 6 are rejected under 35 USC §103(a) as being unpatentable over Logan, in view of Yahagi, and claims 17 and 20 are rejected under 35 USC §103(a) as being unpatentable over Logan in view of Yahagi, and further in view of

5

Ladd. Regarding claims 3 and 17, the Examiner asserts that Logan discloses wherein the step of determining a first time comprises: determining an offset time (start and ending offset) between the time stamp and a time when the first voice command (audio presentation) is active (column 4, lines 48-57); and determining (identify) the first time

5 through reference to the time stamp and the offset time (column 7, lines 41-45). Regarding claim 4, the Examiner asserts that Logan discloses determining an offset time (start and ending offset) between the time stamp and a time when the first voice command (audio presentation) is spoken.

In the text cited by the Examiner, Logan teaches that,

10   as hereinafter described in connection with FIG. 5, each
voice or text program segment preferably includes a sequencing file which
contains the identification of highlighted passages and hypertext anchors
within the program content. This sequencing file may further contain
references to image files and the start and ending offset locations in the
15 audio presentation when each image display should begin and end. In this
way, the image presentation may be synchronized with the audio
programming to provide coherent multimedia programming.
(Col. 4, lines 48-57.)

20   Appellants note that Logan does *not* disclose or suggest "voice commands," as would be apparent to a person of ordinary skill in the art. Claims 3 and 17 require wherein the step of determining a first time comprises: determining an offset time between the time stamp and *a time when the first voice command is spoken*; and determining the first time through reference to the time stamp and the offset time. Claims

25 4 and 20 require determining an offset time between the time stamp and a time when the *first voice command is spoken.* Claim 6 requires determining a first offset time between the first time stamp and a time when the *first voice command is spoken*.

Thus, Logan, Yahagi, and Ladd, alone or in any combination, do not disclose or suggest wherein the step of determining a first time comprises: determining an

30 offset time between the time stamp and a time when the first voice command is spoken; and determining the first time through reference to the time stamp and the offset time, as required by claims 3 and 17, do not disclose or suggest determining an offset time between the time stamp and a time when the first voice command is spoken, as required

by claims 4 and 20, and do not disclose or suggest determining a first offset time between the first time stamp and a time when the *first voice command is spoken*, as required by claim 6.

<u>Claims 13 and 14</u>

Claims 13 and 14 are rejected under 35 USC §103(a) as being unpatentable over Logan in view of Yahagi, and further in view of Ladd. Regarding claim 13, the Examiner asserts that Logan discloses packaging the first time, second time (total time), and one task name (plays field) from the at least one task name into a time increment (col. 19, line 63, to col. 20, line 7).

Appellants note that, in the text cited by the Examiner, Logan teaches that

> the Duration field of the Program_Segment record specifies the duration of the program segment expressed in seconds. The Plays field is an accumulator field which is incremented by incoming Usage_Log records to reflect the total number of times a given program segment has been actually played by all subscribers. Similarly, the TotalTime value represents the total time a given program segment has been actually played by users. Together, these records can be used to determine the advertising fee due from the advertiser, or royalty amount payable to the content provider (the advertiser or content provider being specified the ProviderID field) for the use of this segment.
> (Col. 19, line 63, to col. 20, line 7.)

Appellants could find no disclosure or suggestion by Logan of *packaging the first time, second time, and one task name from the at least one task name into a time increment, or of creating a plurality of time increments, each time increment comprising two times of the first time, second time, or additional plurality of voice command times and a task name.* Claim 13 requires packaging the first time, second time, and one task name from the at least one task name into a time increment. Claim 14 requires determining a plurality of task names from the text; associating a task name with two of the first time, second time, or additional plurality of voice command times; creating a plurality of time increments, each time increment comprising two times of the first time, second time, or additional plurality of voice command times and a task name.

Thus, Logan, Yahagi, and Ladd, alone or in any combination, do not disclose or suggest packaging the first time, second time, and one task name from the at

7

least one task name into a time increment, as required by claim 13, and do not disclose or suggest determining a plurality of task names from the text; associating a task name with two of the first time, second time, or additional plurality of voice command times; creating a plurality of time increments, each time increment comprising two times of the

5   first time, second time, or additional plurality of voice command times and a task name, as required by claim 14.
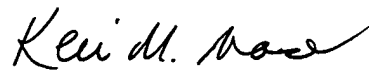

Conclusion

The rejections of the cited claims under section 103 in view of Logan,

10   Yahagi, and Ladd, alone or in any combination, are therefore believed to be improper and should be withdrawn. The remaining rejected dependent claims are believed allowable for at least the reasons identified above with respect to the independent claims.

The attention of the Examiner and the Appeal Board to this matter is appreciated.

15   Respectfully,

                                                        Kevin M. Mason
Date: February 10, 2006   Kevin M. Mason
20   Attorney for Applicant(s)
                                                        Reg. No. 36,597
                                                        Ryan, Mason & Lewis, LLP
                                                        1300 Post Road, Suite 205
                                                        Fairfield, CT 06824
25   (203) 255-6560

APPENDIX

1.   A method, performed on a computer system, for tracking time using speech recognition, the method comprising the steps of:

5   accessing speech data;

              recognizing at least two voice commands from the speech data, each voice command occurring at a different time;

              determining a first time associated with a speaking of a first of the voice commands, wherein said first voice command identifies a start of a time interval;

10   determining a second time associated with a speaking of a second of the voice commands, wherein said second voice command identifies an end of said time interval; and

              storing data identifying said time interval and data identifying one or more of said first voice command and second voice command.

15

2.   The method of claim 1, wherein the second voice command is implied because a predetermined time from the first voice command elapses before another voice command occurs or because a predetermined ending time occurs and there is no voice command after the first voice command but before the predetermined ending time, and

20   wherein the step of determining a second time comprises the step of assigning the second time as the predetermined time plus the first time, if the first voice command elapses before another voice command occurs, or as the predetermined ending time, if the predetermined ending time occurs and there is no voice command after the first voice command but before the predetermined ending time.

25

3.   The method of claim 1, wherein:

              the speech data comprises a time stamp;

              the step of determining a first time comprises:

                      determining an offset time between the time stamp and a

30   time when the first voice command is spoken; and

                      determining the first time through reference to the time

stamp and the offset time.

4.   The method of claim 1, wherein:

the speech data comprises a time stamp;

5   the step of determining a first time comprises:

determining an offset time between the time stamp and a time when the first voice command is spoken; and

determining the first time through reference to the time stamp and the offset time; and

10   the step of determining a second time comprises:

determining a second offset time between the time stamp and a time when the second voice command is spoken; and

determining the second time through reference to the time stamp and the second offset time.

15

5.   The method of claim 4, wherein:

the step of determining the first time through reference to the time stamp and the offset time comprises the step of adding the offset time to the time stamp to determine the first time; and

20   the step of determining the second time through reference to the time stamp and the second offset time comprises the step of adding the second offset time to the time stamp to determine the second time.

6.   The method of claim 1, wherein:

25   the speech data comprises first and second time stamps;

the step of determining a first time comprises:

determining a first offset time between the first time stamp and a time when the first voice command is spoken; and

determining the first time through reference to the first time

30   stamp and the first offset time; and

the step of determining a second time comprises:

  determining a second offset time between the second time stamp and a time when the second voice command is spoken; and

  determining the second time through reference to the

5 second time stamp and the second offset time.

7. The method of claim 1, further comprising the steps of:

recording speech onto a portable recorder; and

loading the speech data from the portable recorder to the computer system,

10 the speech data comprising the speech and a plurality of time stamps.

8. The method of claim 1, further comprising the step of:

determining at least one task name from the text of the at least two voice commands.

15

9. (Canceled)

10. The method of claim 8, wherein the step of determining at least one task name comprises finding the at least one task name in the text.

20

11. The method of claim 8, wherein the step of determining at least one task name comprises associating at least one task name to said time interval between the first and second times, wherein the at least one task name is not in the text.

25 12. The method of claim 8, wherein the at least one task name comprises two task names, a first task name associated with a first of the voice commands and a second task name associated with a second of the voice commands, wherein the first and second voice commands occur adjacent to each other in time, wherein the first and second task name are different, and wherein the second voice command is assumed to end a first task corresponding to the first task name and start a second task corresponding to the second

30

task name.

13.   The method of claim 8, further comprising the step of packaging the first time, second time, and one task name from the at least one task name into a time

5   increment.

14.   The method of claim 8, wherein the at least two voice commands comprises a plurality of voice commands, wherein the at least one task name comprises a plurality of task names, and wherein the method further comprises the steps of:

10   determining an additional plurality of voice command times, each of the voice command times associated with one of the plurality of additional voice command times;

        converting each of the plurality of voice commands to text;

        determining a plurality of task names from the text;

15   associating a task name with two of the first time, second time, or additional plurality of voice command times;

        creating a plurality of time increments, each time increment comprising two times of the first time, second time, or additional plurality of voice command times and a task name; and

20   storing the plurality of time increments.

15.   A method, performed on a computer system, for tracking time using speech recognition, the method comprising the steps of:

        accessing speech data comprising a plurality of time stamps and speech;

25   converting the speech to text;

        composing a plurality of voice commands from words in the text, each voice command corresponding to a phrase grammar rule;

        determining a time associated with a speaking of each of the voice commands;

30   determining a plurality of tasks, each task associated with at least one of

the times and at least one of the voice commands;

determining a plurality of time increments, each time increment comprising one of the tasks and at least one of the times; and

storing one or more of said time increments.

16.   The method of claim 15, wherein the step of determining a plurality of tasks comprises determining a task name for each of the plurality of tasks, wherein first and second voice commands occur adjacent to each other in time, wherein a first task name is associated with the first voice command and a second task name is associated with the second command, wherein the first and second task name are different, and wherein the second voice command is assumed to end a first task corresponding to the first task name and start a second task corresponding to the second task name.

17.   The method of claim 15, wherein:

the speech data comprises a first time stamp;

the step of determining a time comprises:

determining a first offset time between the first time stamp and a time when the first voice command is spoken; and

determining the first time through reference to the first time stamp and the first offset time.

18.   The method of claim 15, further comprising the steps of:

recording speech onto a portable recorder; and

loading the speech data from the portable recorder to the computer system, the speech data comprising the speech and the plurality of time stamps.

19.   A system for tracking time using speech recognition, the system comprising:

a computer system comprising:

a memory that stores computer-readable code; and

a processor operatively coupled to the memory, the processor configured to implement the computer-readable code, the computer-readable code configured to:

access speech data;

recognize at least two voice commands from the speech data, each voice command occurring at a different time;

determine a first time associated with a speaking of a first of the voice commands, wherein said first voice command identifies a start of a time interval;

determine a second time associated with a speaking of a second of the voice commands, wherein said second voice command identifies an end of a time interval;

convert each of the at least two voice commands to text;

determine text versions of the at least two voice commands by comparing words in the text with phrase grammar rules; and

storing data identifying said time interval and data identifying one or more of said first voice command and second voice command.

20.   The system of claim 19, wherein the speech data comprises a time stamp, and wherein the computer-readable code is further configured to:

when determining a first time:

determining an offset time between the time stamp and a time when the first voice command is spoken; and

determining the first time through reference to the time stamp and the offset time; and

when determining a second time:

determining a second offset time between the time stamp and a time when the second voice command is spoken; and

determining the second time through reference to the time stamp and the second offset time.

21.   The system of claim 19, wherein the computer-readable code is further configured to:

          store the time increments; and

          place the time increments into a file having a format suitable for importing into a time and billing program.

22.   The system of claim 19, wherein the system further comprises a digital personal recorder and wherein the computer-readable code is further configured to receive the speech data from the digital personal recorder.

23.   (Canceled)

24.   A system for tracking time using speech recognition, the system comprising:

          a computer system comprising:

          a memory that stores computer-readable code; and

          a processor operatively coupled to the memory, the processor configured to implement the computer-readable code, the computer-readable code configured to:

          access speech data comprising a plurality of time stamps and speech;

          convert the speech to text;

          compose a plurality of voice commands from words in the text, each voice command corresponding to a phrase grammar rule;

          determine a time associated with a speaking of each of the voice commands;

          determine a plurality of tasks, each task associated with at least one of the times and at least one of the voice commands;

          determine a plurality of time increments, each time increment comprising one of the tasks and at least one of the times; and

          storing one or more of said time increments.

25.   The system of claim 24, wherein the computer-readable code is further configured to:

place the time increments; and

place the time increments into a file having a format suitable for importing

5   into a time and billing program.

26.   The system of claim 24, wherein the system further comprises a digital personal recorder and wherein the computer-readable code is further configured to receive the speech data from the digital personal recorder.

10

27.   An article of manufacture comprising:

a computer readable medium having computer readable code means embodied thereon, the computer readable program code means comprising:

a step to access speech data;

15   a step to recognize at least two voice commands from the speech data, each voice command occurring at a different time;

a step to determine a first time associated with a speaking of a first of the voice commands, wherein said first voice command identifies a start of a time interval;

a step to determine a second time associated with a speaking of a second

20   of the voice commands, wherein said second voice command identifies an end of a time interval;

a step to convert each of the at least two voice commands to text;

a step to determine text versions of the at least two voice commands by comparing words in the text with phrase grammar rules; and

25   storing data identifying said time interval and data identifying one or more of said first voice command and second voice command.

28.   An article of manufacture comprising:

a computer readable medium having computer readable code means

30   embodied thereon, the computer readable program code means comprising:

16

a step to access speech data comprising a plurality of time stamps and speech;

a step to convert the speech to text;

a step to compose a plurality of voice commands from words in the text,

5   each voice command corresponding to a phrase grammar rule;

a step to determine a time associated with a speaking of each of the voice commands;

a step to determine a plurality of tasks, each task associated with at least one of the times and at least one of the voice commands; and

10   a step to determine a plurality of time increments, each time increment comprising one of the tasks and at least one of the times; and

storing one or more of said time increments.

29.   The method of claim 8, further comprising the step of determining text

15   versions of the at least two voice commands by comparing words in the text with phrase grammar rules.

30.   The system of claim 19, wherein the computer-readable code is further configured to:

20   convert each of the at least two voice commands to text; and

determine text versions of the at least two voice commands by comparing words in the text with phrase grammar rules.

EVIDENCE APPENDIX

There is no evidence submitted pursuant to § 1.130, 1.131, or 1.132 or entered by the Examiner and relied upon by appellant.

RELATED PROCEEDINGS APPENDIX

There are no known decisions rendered by a court or the Board in any proceeding identified pursuant to paragraph (c)(1)(ii) of 37 CFR 41.37.